



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification<sup>6</sup>:</b> <b>C12Q 1/68, C12P 19/34, C07K 13/00,</b> <b>G06F 17/00</b>	<b>A1</b>	<b>(11) International Publication Number:</b> <b>WO 00/05414</b> <b>(43) International Publication Date:</b> 3 February 2000 (03.02.00)
<b>(21) International Application Number:</b> PCT/US99/16417 <b>(22) International Filing Date:</b> 21 July 1999 (21.07.99)  <b>(30) Priority Data:</b> 60/093,641 21 July 1998 (21.07.98) US 09/181,601 29 October 1998 (29.10.98) US  <b>(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications</b> US 60/093,641 (CIP) Filed on 21 July 1998 (21.07.98) US 09/181,601 (CIP) Filed on 29 October 1998 (29.10.98)  <b>(71) Applicant (for all designated States except US):</b> RUTGERS, THE STATE UNIVERSITY OF NEW JERSEY [US/US]; ASB, Annex 11, 58 Bevier Road, Piscataway, NJ 08854-8010 (US).  <b>(72) Inventors; and</b> <b>(75) Inventors/Applicants (for US only):</b> ANDERSON, Stephen [US/US]; 158 Springdale Road, Princeton, NJ 08540 (US). MONTELLONE, Gaetano [US/US]; 127 North 5th Avenue, Highland Park, NJ 08904 (US). HUANG, Yuanpeng		[US/US]; ASB, Annex 11, 58 Bevier Road, Piscataway, NJ 08854-8010 (US).  <b>(74) Agents:</b> AUERBACH, Jeffrey, I. et al.; Howrey & Simon, 1299 Pennsylvania Avenue, N.W., Box 34, Washington, DC 20004-2402 (US).  <b>(81) Designated States:</b> AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).  <b>Published</b> <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>
<b>(54) Title:</b> LINKING GENE SEQUENCE TO GENE FUNCTION BY THREE DIMENSIONAL (3D) PROTEIN STRUCTURE DETERMINATION		
<b>(57) Abstract</b> <p>The present invention provides a structure-functional analysis engine for the high-throughput determination of the biochemical function of protein domains of unknown function, as exemplified in the flowchart of the figure. The present invention uses bioinformatics, molecular biology and nuclear magnetic resonance tools for the rapid and automated determination of the three-dimensional structures of proteins and protein domains.</p> <div data-bbox="862 1108 1448 1948"> <p style="text-align: center;"><b>NOESY ASSIGN PROCESS</b></p> <pre> graph TD     S1[STEP 1: SIMPLE MATCH-GET ALL POSSIBLE ASSIGNMENT (A-TYPE MATCH).] --&gt; S2[STEP 2: IDENTIFY EXPECTED PEAKS WHICH ARE INTRA/SED. OR CONSISTENT WITH SECONDARY STRUCTURE. PUT IN U AND E SET. CREATE CLUST.]     S2 --&gt; S3[STEP 3: DEFINE LOCAL MATCH TOLERANCE FOR HX DIMENSION BASED ON ASSIGNED HX RESONANCES FROM E SET.]     S3 --&gt; S4[STEP 4: SUPPLEMENT U BASED ON CHEMICAL SHIFT (UNAMBIGUOUS) WITH NOISE FILTER.]     S4 --&gt; S5[STEP 5: DEFINE B-TYPE MATCHES OF A-TYPE MATCHES.]     S5 --&gt; S6[STEP 6: USE CF OF E TO ASSIGN B TO U.]     S6 --&gt; S7[STEP 7: USE STM TO ASSIGN B TO U.]     S7 --&gt; S8[STEP 8: USE HP-CORE TO ASSIGN B TO U.]     S8 --&gt; S9[STEP 9: COMPUTE 3D STRUCTURE.]     S9 --&gt; S10[STEP 10: VALIDATION-REMOVE FROM U LIST THE CONSISTENTLY VIOLATED NOE ASSIGNMENTS.]     S10 --&gt; S11[STEP 11: IDENTIFY EXPECTED PEAKS THAT ARE CONSISTENT WITH 3D STRUCTURE. PUT IN U SET.]     S11 --&gt; S12[STEP 12: BACK CALCULATE ALL POSSIBLE NOE'S THAT ARE EXPECTED FROM STRUCTURE. OUTPUT ANY PREDICTED ASSIGNMENTS NOT IN U OR A LIST AND PEAKS STILL IN THE A LIST.]           </pre> </div>		